

HT1220 Report

RECORD INFORMATION

Gene ID: 1220
Sequence ID: 1220
Protein ID: 1220
Sequence name: thrombospondin 1, alt. transcript 1
Genome: nucleus
Taxon: Homo sapiens
Locus: 1220
Common Name: thrombospondin 1
Role ID: 40

Coding sequence length: 3513 nt
Transcript sequence length: 5722 nt
Expression data: THC201673

ACCESSION DATA

HT1220 is derived from accessions(s):

SP:PO7996 (THROMBOSPONDIN 1 PRECURSOR.)
GB:X04665 (Human mRNA for thrombospondin)
GB:X14787 (Human mRNA for thrombospondin)
GB:U12471 (thrombospondin-p50 (Homo sapiens))
GB:M99425 (Human thrombospondin mRNA, 3' end.)
PIR:G01478 (thrombospondin-p50 - human (fragment))
GB:U12471 (Human thrombospondin-1 gene, partial cds.)
GB:J04835 (Human thrombospondin gene, exons 1, 2 and 3.)
GB:M25631 (Homo sapiens (clone lambda-TS-33) thrombospondin (THBS) mRNA, 5' end.)

ALTERNATIVE SPLICING INFORMATION

Alternative splice forms for this gene:

HT3987 thrombospondin 1, alt. transcript 2

MAPPING DATA

GDB accession(s) for this gene:

GDB ID: Symbol

FIGURE 1A

gdb:120438 THBS1

cDNA FEATURES

Feature	End 5	End 3
coding_seq	112	3624
3'UT	3625	5722
spjunc_h	1235	1236

SEQUENCE

nucleotide:

ggacgcacaggcattccccggccccctccagccccctggccgcctcgccaccgctccccggc
cgccgcgctccggtaacacacaggatccccgtctgggaccaacagctccaccatggggctg
gcctggggactaggcgctctgttctgtatgcattgtgtggaccaaccgcattccagag
tctggggagacaacagcgtttgacatcttgactcacccgggcccggcccaagggg
tctggggccgacttggtaaggggccccgacccttccagcccagcttccgcattcgaggat
gccaacctgateccccctgtgcctgatgacaagttcaagacccctggatgtgtgcgg
gcagaaaaagggtttctcttctggcattccctgaggcagatgaagaagacccggggcactc
ctgctggccctggagcggaaagaccacttgcaggtttcagcgtgggtccaaatggc
aaggcgggaccctggacctcagcgtgaccgtccaaaggaaagcagcagcgtgggtctgt
gaagaagctctctggcaaccggccagtgaaagagacatcacccctttgtcaggaagac
agggcccagctgtacatcgactgtaaaaagatggagaatgtcgtgagttggacgtccccatc
caaagcgtttcaccagagacttgcagcatgcctgactccgcattcgcggaaagggg
gtcaatgacaatttccaggggggtgtcagaaatgtgaggtttgcataaccacacca
gaagacatcttcaggaacaaaaggctgtccagcttaccagtgtcctctccatcccttgc
aacaacgtgtgaatggtccagccatccgcactaactacatggccacaagaca
aaggacttgcagccatctggggcatctctgtatgagctgtccagcatggtcttggaa
ctcagggcctgcgaccattgtgaccacgctgcaggacagcatccgcggaaatgtactgaa
gagaacaaagagtggccaatgagctgaggcggccctccatgtctatcacaacggagtt
cagtagaaaataacgaggaatggactgttgcatactgcactgagttgcactgtcagaac
tcagttaccatctgcaaaaagggttcttgccttgccttgcggccatgtccacagtt
cctgatggagaatgtgtcctcgctgttggccctccagcactctgcggacgtggctgt
ccatggccggacttgcctgttgcatacgactgtggcaatggaaattcagcagcggcc
cgctccgtcgatagcctcaacaccgatgtgagggttgcctcggtccagacacggacatgc
cacattcaggagtgtgacaaaagatttaaacaggatgggtggctggagccactgtcccc
tggtcatttgcattgtgacatgtgggtgtatgggtgcatacaggatccggctctgcaac
tctccagccccccagatgaaatggaaaaccctgtgaaggcgaagcggggagaccaaaagcc
tgcagaaaaagacgcctgccccatcaatggaggctgggtcttgcattgtcaccatgggacatc
tgcattgtcaccctgtggaggagggtacagaaaacgttagtgcattgtcaacaaaccggca
ccccagtttggaggcagggactgcgttgcattgtaaacagaaaaaccaggatgtcaacaaag
caggactgtccaaattgtggatggatgcctgttgcattgttgcggggatgtact
agtcaccctgtatggcagctgaaatgtgggtgttgccttgcattgtggaaaatggc
atccagtgcacagatgttgcattgtggatgtcaacaaaggatgtcattgtatggggat
ggagggcaccgggtgtgagaaacacggacccggctacaactgcattgcctggcc
ttcaccggctcacagcccccttgcggccagggtgtcgaacatgcacggccaaacaaacagggt
tgcangggccgtacccctgcacggatgggacccacgactgcacacaagaacggccaaatgt
aactacccggccactatagcgcacccatgtaccgcgtgcagttgcacgcattgcct
ggcaatggcatcatctgcggggaggacacagacccctggatgggtggcccaatgagaac
gtgtgcgtggccaaatgcgacttaccactgcacaaaaggataatgccttgcatt
tcaggggcaggaagactatgacaaggatggaaattgggtgtgcctgtatgtac
aatgataaaattccagatgacaggacacactgtcattccattacaacccagctcag
gactatgacagagatgtgtggagacccgcgtgtgacaactgtccctacaacc
acaacccaa

FIGURE 1B

gacgggtatctcaatgaacggacaactgccagtcgttacaatgtggaccagagagac
actgatatggatggggtggagatcagtgtgacaattgccttggaaacacaatccggat
cagctgactctgactcagaccgcattggagatacctgtgacaacaatcaggatattgt
gaagatggccaccagaacaatctggacaactgtccctatgtgccaatgccaaccaggct
gaccatgacaagatggcaagggagatgcctgtgaccacgatgtgacaacgatggcatt
cctgtatgacaaggacaactgcagactcgtgccaatccgaccagaaggactctgacggc
gatggtcaggtgtgcctgcaagatgatttgaccatgacagtgtgccaagacatcgat
gacatetgtcctgagaatgtgacatcagtgagaccgatttccggattccagatgatt
cctctggacccaaagggacatccaaaatgaccctaactgggttgcacggccatcagggt
aaagaactcgtccagactgtcaactgtgatcctgactcgctgttaggttatgtgagtt
aatgetgtggacttcagtggcaccttcttcatcaacaccgaaagggacatgactatgt
ggatttgtcttgctaccagtcagcagccgtttatgttgtgatgtggaaagcaagtc
acccagtcctactggacacccaacccacgagggctcagggatactcgggccttctgtg
aaagttgtaaactccaccacaggcctggcgagcacctgccaacgcctgtggcacaca
ggaaacacccctggccagggtgcgcacccgtggcatgaccctgtcacataggctggaa
gatccacgcctacagatggcgttcagccacaggccaaagacgggttcatagatgt
gtgatgtatgaagggaaagaaaatcatgctgactcaggaccatctatgataaaacctat
gctgggttagacttagggtttgtcttcataagaaatgggtttctctgacctg
aaatacaaatgttagagatccctaatcatcaaattgtgatgtggaaagactgatcataaacc
aatgctgtattgcaccccttggactatgggttgcagaaaacccaggatcacttctt
cttggcttccttccttcgtgctgcatcagtgtggactctagaacgtgcacccgtcc
tcaagaaaatgcagtttcaaaaacagactcatcagcattcagccatgtgatgaaagac
atcttccaagcatataaaatgtcttgggttctttgaaaagcatctacttgc
agttggaaagggtgcccattccactctgccttgcacagagcagggtgttgcattgtgaggg
catctcgagcagtgactcaaaagcatttcagcatgtcagagaagggaggactca
agaattagcaacaaaaccaccctgacatcccttcaggacacacggggagcagaggca
aagcactaaggggaggcgcatacccgagacgatgtatgtggaaatatggaggaactg
ttacatgttcggactaagtcatttcagggttgcagaaaagactattgtggatccatgt
tgctgactggcgttagctgatattttcagggttgcagaaaagactattgtggatccatgt
ggagacaaagactggcttctggacttctccctgatccccacccttactcatcacttgc
agtggcagaattaggaaatcagaatcaaaccagtgtatggcactgtgtggatccattgc
ctggtcacattgaaattgggttgcatttcatttagatgttagtgcagatgtatgtgg
aataggaaaacccatctcgttgcacaaaattttatcaacctaactaaaacatcc
ttctcttttcgttaattacttaggttagtttcttgcacaaaattttatcaacctaactaaaacatcc
ttttaaagtcttacgtatgtaaaatattttacttattctggaaagatctggctg
aggattattcatggAACAGGAAGAGCTAAAGACTATCCATGTCTTGTGAGAGT
CTTCGTACTGTAAGATTGTAATACAGATTATTATTAACTCTGTCTGCTGGAAATT
TAGGCTTCATACGGAAAGTGTGAGAGCAAGTAGTGTGACATTATCAGCAAATCTTG
CAAGAACAGCACAAGGAAATCAGTCATAATAAGCTGTCTGCCCTTGTGCTCAGAGTGG
ATGTTATGGATTCTTCTGTGTTATCTTCAAGTGGAAATTAGTGGTTATCC
ATTGCAATGTTAAATTGCAAAGAAAGCCATGAGGTCTCAATACTGTGTTACCCCA
TCCCTGTGCTATTCAGGGAGAAGGAAAGCATATAACCTTTCTTCAATT
AAAAGAGAAAATGACAAAAGGTGAAACTACATACAAATATTACCTCATTTGTG
TGACTGAGTAAAGAATTGGATCAAGCGGAAAGAGTAAAGTGTCTAACAAACCTAA
GCTACTGTGACTACCTAAAAGTCAGTGTGACATAGCATAAAACTCTGAGAGAAGT
TTCCCAATAAGGAAATAGCATTGAAATGTTAAATACAAATTCTGAAAGTTATGTT
CTCATCTGGTATACCATGCTTATTTATAAATTATTCTCATTCAGGCAATTGG
TAGAATATTCAAGATTGTGAGATATGCTATTAAATAATTCTGACACATTGAAAG
AGAGTTAGTATTCTATTCTATATAATGTTGACACTGAATTGAAAGATTGTT
TTCTTTTTGTTTTTTTTTTTTTTTTGTTGTTGACCTCCATT
CTATTGCAATAACCTTTCTAGGAATGTCGTTTTTTGTCACACATTGTTATCCATT
CACATTCTAAAGCAGTGTAGTGTATATTACTGTTCTTGTACAAGGAACAACA
AAATCATATGGAAATTATTT

protein:

MGLANGLGVLFLMHVCGTNRIPESGGDNSVFDIFELTGAAKGSORRLVKGPDPSSPAFR

FIGURE 1C

IEDANLIPPVPPDDKFQDLVDAVRAEKGFLLASLRQMKKTRGTLLALERKDHSGQVFSVV
SNGKAGTLQLSLTVQGKQHVSVVEEALLATGQWKSITLFVQEDRAQLYIDCEKMENAELD
VPIQSVFTRDLASIARLRIAKGGVNDNFQGVLQNVRFVFGTTPEDILRNKGCSSTSVELL
TLQNNVVNGSSPAIRTNIGHKTKDLQAIKGISCDELSMVLELRGLRTIVTLQDSIRK
VTBENKELANELRRPLCYHNGVQYRNNEWTVDSCTECHCQNSVTICKVSCPIMPSCN
ATVPDGECCPRCWPSSDADDGWPSPWSEWTSCTSCCGNIQQQRGRSCDSLNNRCEGSSVQT
RTCHIQECDKRFKQDGGWSHSPWSSCSVTCDGVIIRLCLNSPSPQMNKGKCEGEARE
TKACKKDACPINGGPWSPWDICSVTCGGVQKRSRLCENNPAQFGGKDCVGDVTEENQI
CNKQDCPIDGCLSNPCFAGVKCTSYPDGSWKCGACPGYSGNGIQCTDVECKEVPDACP
NHNGEHRCENTDPGYNCLPCPPRFTGSQPGQGVEHATANKQVCKPRNPCTDGTHDCNKN
AKCNYLGHYSDPMYRCECKPGYAGNGIICGEDTLDGWPNENLVCVANATYHCKKUNCPN
LPNSGQEDYDKDGIGDACDDDDNDKIPDDRDNCPHYNPAQYDYDRDDVGDRCDNCPYN
HNPDQADTDNNNGEDACAADIDGDGILNERDNCQYVNVDQRDTDMGVDQCDNCPLRH
NPDQLDSDSDRIGDTCNNQDIDEDGHQNNLDNCPVVPNANQADHDKGKDACDHDDDN
DGPDDKDNCRLVPNPDQKDSGDGRGACKDDFDHDSVPDIDICPENVDISETDFRRF
QMIPLDPKGTSQNDPNWVVRHQGKELVQTVNCDPGLAVGYDEFNAVDFSGTFFINTERDD
DYAGFVFGYQSSSRFYVVMWKQVTQSYWDTNPTRAQGYSGLSVKVNNTTGPGEHLRNAL
WHTGNTPGQVRTLWHDPRHIGWKDFTAYRWRLSHRPKTGFIRVVMYEGKKIMADSGPIYD
KTYAGGRRLGLFVFSQEMVFFSDLKYECRDP

FIGURE 1D

SEQUENCE- TSP-2

nucleotide:

acggcatcca gtacagaggg gctggacttg gaccctgca gcagccctgc
acaggagaag cgccatataa agccgcgctg cccgggagcc gctcggccac
gtccaccgga gcacccctgca ctgcagggcc ggtctctcgc tccagcagag
cctgcgcctt tctgactcgg tccggAACAC tggaaaccagt catcaactgca
tcttttggc aaaccaggag ctcagctgca ggaggcagga tggctctggag
gctggcctg ctggctctgt gggtgtggcc cagcacccaa gctggtcacc
aggacaaaaga cacgacccttc gacctttca gtatcagcaa catcaaccgc
aagaccattg gcgccaagca gttccgcggg cccgaccccg gcgtgcccggc
ttaccgcttc gtgcgctttg actacatccc accggtaac gcagatgacc
tcagcaagat caccaagatc atgcggcaga aggagggctt ctccctcact
gcccaagctca agcaggacgg caagtccagg ggcacgctgt tggctctgg
gggccccggg ctctccaga ggcagttcga gatcgtctcc aacggccccc
cgacacgct gatatcacc tactggattt acggcaccccg gcatgtggc
tccctggagg acgtcggcct ggctgactcg cagtggaaga acgtcaccgt
gcaggtggct ggcgagacct acagttgca cgtggcgtgc gacctcatag
gaccagttgc tctggacgag cccttctacg agcacctgca ggcggaaaag
agccgatgt acgtggccaa aggctctgcc agagagatc acttcagggg
tttgcttcag aacgtccacc tagtgttga aaactctgtg gaagatattc
taagcaagaa gggttgccag caaggccagg gagctgagat caacgccatc
agtgagaaca cagagacgct ggcctgggt cccatgtca ccaccgagta
cgtggccccc agctcgaga ggaggccgaa ggtgtgcgaa cgctcggtcg
aggagctggg aaacatggc caggagctt cggggctcca cgtcctcggt
aaccagctca gcgagaacct caagagatg tcaaatgata accagtttct
ctggagactc attggggcc ctcctaagac aaggaacatg tcaagttgt
ggcaggatgg ccggttctt gcgaaaatg aaacgtgggt ggtggacagc
tgcaccacgt gtacctgcaaa gaaattttaa accatttgcc accaaatcac
ctgcccgcct gcaacctgca ccagtccatc ctttggaa ggcgaatgct
gcccttcctg cttccactcg gtggacgggtg aggagggctg gtctccgtgg
gcagagtgg a ccagtgcct cgtacgtgt ggctctggaa cccagcagag
aggccggcct tgtgacgtca ccagcaacac ctgctgggg ccctcgatcc
agacacgggc ttgcagtctg agcaagtgtg acacccgcatt cggcaggac
ggcggctgg a gccactggc accttggctc tcatgctctg tgacctgtgg
agttggcaat atcacacgc tccgtctctg caactccca gtccccaga
tggggggcaaa gaattgcaaa gggagtggcc gggagaccaa agcctgccc
ggcgcggccat gccaatcgat tggccgtgg agccccctggt ccccggtgg
ggcctgcact gtcacccgtt ccgggtggat cccggagcgc acccggtct
gcaacagcccc tgagccctcag tacggaggaa aggccgtcgat gggggatgt
caggagcgtc agatgtgcaaa caagaggagc tgccccgtgg atggctgttt
atccaacccc tgcttccgg gagcccaatgt cagcagctt cccgatgggt
cctggcgtatg cggcttctgc cctgtgggt tcttggcaaa tggcaccac
tgtgaggacc tggacgtgt tgccctggc cccgacatct gcttctccac
cagcaaggtg cctcgctgt tcaacactca gcctggctt cactgcctgc
cctggccgcct ccgatacaga gggaaaccagc ccgtcggtt cggcctggaa
gcagccaaga cggaaaagca agtgtgtgag cccgaaaacc catgcaagga
caagacacac aactgccaca agcacgcgaa gtgcacatctac ctgggtcact
tcagcgaccc catgtacaag tgcgagtgcc agacaggcta cggggcgcac
gggctcatct gcgggagga ctcggacctg gacggctggc ccaacctcaa
tctggctctgc gccaccaacg ccacctacca ctgcacatcaag gataactgcc

FIGURE 2A

ccccatctgcc aaattctggg caggaagact ttgacaagga cgggattggc
gatgcctgtg atgatgacga tgacaatgac ggtgtacccg atgagaagga
caactgccag ctcctcttca atccccgcca ggctgactat gacaaggatg
aggttgggaa ccgctgtgac aactgccctt acgtgcacaa ccctgcccag
atcgacacag acaacaatgg agagggtgac gcctgctccg tggacattga
tggggacgat gtcttcaatg aacgagacaa ttgtccctac gtctacaaca
ctgaccagag ggacacggat ggtgacgggtg tggggatca ctgtgacaac
tgcaccctgg tgcacaaccc tgaccagacc gacgtggaca atgacattgt
tgyggaccag tgtgacaaca acgaggacat agatgacgac ggccaccaga
acaaccagga caactgcccc tacatctcca acgccaacca ggctgaccat
gacagagacg gccagggcga cgcctgtgac cctgatgatg acaacgatgg
cgtccccgat gacagggaca actgcccggct tggatcaac ccagaccagg
aggacttggaa cgggtgata tttgtaaaga tgatttgac
aatgacaaca tcccagatat tggatgtg tggctctaaa acaatgccc
cagtgagaca gacttcagga acttccagat ggtcccccgg gatcccaaag
ggaccaccca aattgatccc aactgggtca ttgcctcatca aggcaaggag
ctggttcaga cagccaaactc ggaccccgcc atcgctgtag gtttgacga
gtttgggtct gtggacttca gtggcacatt ctacgtaaac actgaccggg
acgacgacta tgctggcttc gtctttgggtt accagtcaag cagccgcttc
tatgtggta tatggctact ccggcgtgtc cctcaaggtg gtgaactcca
cacgcgggccc gggcgagcac ctgaggaacg cgctgtggca cacggggAAC
ccacggggac aggtgcgaac cttatggcac gaccccgaga acattggctg
acgccccggc gaaggactac acggctata ggtggcacct gactcacagg cccaagaccg
gctacatcag agtcttagtg catgaaggaa aacaggtcat ggcagactca
ggacctatct atgaccaaac ctacgctggc gggcggctgg gtctatttgt
cttctctcaa gaaatggtct atttctcaga cctcaagtagc gaatgcagag
atatttaaac aagatttgct gcatttccgg caatgcctg tgcatgccc
ggtccttaga cacctcagtt cattgtggtc ttgcggctt ctctctctag
cagcacctcc tggcccttga ccttaactct gatggttctt cacctcttc
cagcaacccc aaacccaaat gccttcagag gataaatatc aatggaaactc
agagatgaac atctaaccca ctagagaaaa ccagttgggt gatatatgag
actttatgtg gagtggaaaat tgggcatgcc attacattgc ttttttttgt
ttgtttaaaaa agaatgacgt ttacatataa aatgtaatta ctatttgtat
ttatgtgtat atggagttga agggaaatact gtgcataagc cattatgata
aattaagcat gaaaaatatt gctgaactac ttttgggtct taaagttgtc
actattcttgc aatttagagtt gctctacaat gacacacaaa tcccgctaaa
taaattataa acaagggtca attcaaattt gaagtaatgt tttagtaagg
agagattaga agacaacagg catagcaaat gacataagct accgattaaac
taatcggaac atgtaaaaca gttacaaaaa taaacgaact ctcctcttgt
cctacaatga aagccctcat gtgcagtaga gatgcagtt catcaaagaa
caaacatcct tgcaaattgg tggacgcgg ttccagatgt ggatttggca
aaacctcatt taagtaaaaag gttagcagag caaatgtcgg tgcttagct
gctgcttgtg cgggttgtggc gtcggggagg ctcctgcctg agcttccttc
cccagctttg ctgcctgaga ggaaccagag cagacgcaca ggccggaaaa
ggcgcacatcta acgcgtatct aggctttgggt aactgcggac aagttgctt
tacctgattt gatgatacat ttcatggaa ttccagttat aaatattttg
ttaatattta ttaagtgact atagaatgca actccattta ccagtaactt
attttaaaa tgccttagaa cacatatgta gtataatttc tagaaacaaa
catctaataa gtatataatc ctgtggaaaat atgaggcttg ataatatttag
gttgcacga tgaagcatgc tagaagctgt aacagaatac atagagaataa

FIGURE 2B

atgaggagtt tatgatggaa ccttaatata taatgttgcc agcgatttt
gttcaatatt tgtaactgtt atctatctgc ttttatatgga attcttttaa
ttcaaacgct gaaaacgaat cagcatttag tcttgcagg cacacccaaat
aatcagtcat gtgtaatatg cacaagttt tttttgttt ttttttttt
gttgggttgtt tttttgtt taagttcat gatcttctg cagggaaatag
tcactcatcc cactccacat aagggttta gtaagagaag tctgtctgtc
tgatgtgga tagggggcaa atcttttcc ccttctgtt aatagtcac
acatttctat gccaaacagg aacgatccat aacttagtc ttaatgtaca
cattgcatt tgataaaatt aattttgtt tttccttga ggttgcgt
tgtgttgtt tgctgcactt ttactttt tgcgtgtgga gctgtattcc
cgagacaacg aagcgttggg atacttcatt aaatgttagcg actgtcaaca
gcgtgcaggt tttctgttc tttgttgtgg ggtcaaccgt acaatgggt
ggaaatgacg atgatgtgaa tatttagaat gtaccatatt tttttaaat
tatttagtt tttctaaaca aattttatcgat atagggttgc gaaacgtcat
gtgtttgaa aaagactgta aatatttatt tatgtgttca catggtaaaa
atttcaccac tgaaaccctg cacttagcta gaacctcatt tttaaagatt
aacaacagga aataaattgt aaaaaagtt ttct

protein:

MWRLVLLALWWWPSTQAGHQDKDTTFDLFSISNINRKTIGAKQFRGPDPGVPAYRF
VRFDYIPPVNADDLSKITKIMRQKEGFFLTQLKQDGKSRGTLLEGPGLSQRQFE
IVSNGPADTLDLTYWIDGTRHVSLEDVGLADSQWKNTVQAGETYSLHVGCDLIG
PVALDEPFYEHLQAEKSRMYVAKGSARESHFRGLLQNVHLVFENSVEDILSKKGCQQ
GQGAEINAISENTETLRLGPHVTTEYVGPSERRPEVCERSCEELGNMVQELSGLHV
LVNQLSENLKRVSNNDNQFLWELEIGGPKTRNMSACWDGRFAENETWVVDSCTTCT
CKKFKTICHQITCPPATCASPSFVEGECCPSCLHSVGEEGWSPWAETQCSVTCGS
GTQQRGRSCDVTNTCLGPSIQTRACSLSKCDTRIRQDGWSHWSPWSSCSVTGVG
NITRIRLCNSPVPQMGGKNCKGSGRETAKCQGAPCPIDGRWSPWSPWSACTVTCAGG
IRERTRVCNSPEPQYGGKACVGDVQERQMCNKRSCPVDGCLSNPCFPAGAQCSSFPDG
SWSCGFCPVGFLGNHTCEDLDECALVPDICFSTSVPVRCVNTQPGFHCLPCPPRYR
GNQPVGVLGAAKTEKQVCEPENPCKDKTHNCHKHAECIYLGHFSDPMYKCECQTGY
AGDGLICGEDSLDGWPNLNLVCATNATYHC1KDNCPLPNSGQEDFDKDGIIDACD
DDDDNDGVTDEKNCQLLFNPRQADYDKDEVGDRCDCNCPYVHNPAQIDTDNNNEGDA
CSVDIDGDDVFNERDNCPYVYNTDQRDTDGDGVGDHCDNCPLVHNPDQTDVDNDLVG
DQCDNNEDIDDDGHQNNQDNCPYISNANQADHDRDGQGDACDPDDNDGVPDDRDNC
RLVFNPQEDLDGDRGDICKDDFDNDNIPDIIDDVCPENNAISETDFRNQMVPLDP
KGTTQIDPNWVIRHQGKELVQTANSDPGIAVGFDEFGSVDFSGTFYVNTDRDDYAG
FVFGYQSSSRFYVVMWKQVTQTYWEDQPTRAYGYSVSLKVVNSTTGTGEHLRNALW
HTGNTPGQVRTLWHDPRNIGWKDYTAYRWHLTHRPKTGYIRVLVHEGKQVMADSGPI
YDQTYAGGRLGLFVFSQEMVYFSDLKYECRDI

FIGURE 2C

HT2143 Report

RECORD INFORMATION

Gene ID: 2081
Sequence ID: 2143
Protein ID: 2125
Sequence name: thrombospondin 4
Genome: nucleus
Taxon: Homo sapiens
Locus: 2081
Common Name: thrombospondin 4
Role ID: 40

Coding sequence length: 2886 nt
Transcript sequence length: 3074 nt
Expression data: THC168897

ACCESSION DATA

HT2143 is derived from accessions(s):

SP:P35143 (THROMBOSPONDIN 4 PRECURSOR.)
GB:Z19585 (thrombospondin-4 (Homo sapiens))
GB:Z19585 (H.sapiens mRNA for thrombospondin-4)
PIR:A55710 (thrombospondin 4 precursor - human)

cDNA FEATURES

Feature	End 5	End 3
coding_seq	28	2913
3'UT	2914	3074

SEQUENCE

nucleotide:

gaattccggggagcaggaagagccaacatgtggcccgccggagccgcgtcctcctg
ctgcacctggtcctgcagcggtggctagcggcaggcgcccaggccacccccaggcttt
gaccttctccatttccagtccaggctaaaccaggcgctctgtccgtccatgtgaca
gacccegcctgaatgtatcttatgtgatttccaccttcaagctgcagactaaatgtca
gccaccatccggctttactttcaactgacaacagtaaatttgaatttactgtg
atgggacgcgttaagcaaaaggccatccgttacctgaagaacgatggaaagggtgcatttg

FIGURE 3A

1000 2000 3000 4000

gtggtttcaacaacctgcagctggcagacggacaggatccctgaggctg
agcaatttgcagcggggccggctccctagagctcacctggactgcacccaggatgg
tccgttccacaatctcccaaggcccttgcggccctcccaaaaacctgagaccattgaa
ttgaggacttccagaggaagccacaggactcttgcggaaagagctgaagctgggtgaga
ggcteactgttccaggtggccagccctgcaagactgtttcgtcagcagactgacactg
gctgccacaggcacagggacttaacggcagttcttgggtcaaatgacacaattaaac
caactcttggagggatgtgaaaggaccttgcagacagcaggtaaggaaacatcattttg
cgaaacaccatagctgaatgccaggctgcggctctcaagttcagtcctcgacccca
agcacggtggtcgccccggctccccctgcaccccaacacgcccacctcggtgtgac
tcciaaccatgttccaggtgtccatgttgcacactgatggcttccagtgatgg
ccctgccccggatcacacaggaaacgggatcacctgttgcacatgttgcgtgaaa
taccatccctgctacccggcgtgcactgcataaattgtcttgcgttgcacatgtgac
gcctgcccagtggttgcacaggccatgggtcagggtgtggatcagtttgc
tcaaacaaggcggctgcactgcattgtatgtgcgaaatggagcgtgcgttcccaac
tcgatctgcgttaatacttggatcttaccgtgtggccttgcataaggccgggtataact
ggtgatcagataaggggatgcgaaaactgcgaaaacccagagctgaaaccct
tgcagtgtgaatgcccagtgcattgcggatggcggatgtgacatgtgtgtgg
gtcggttggcgtggagatggctatctgtggaaaggatgtggacatgcacatgttaccc
gacgaagaactgcccattgcgtccaggaaactgttgcgaaaatgtgcca
aatttgcgaaaaggatgcgaaaaggatgcgaaaaggatggcattggcgcgttgc
gacggagatggatcctgaatgacggataactgtgttgcattcataatgtggacca
aggaacacgcataaagatatcttgggatgcctgtgataactgcctgagtgtctaaat
aacgaccaggaaaagacaccgtgggatggaaaggagatgcctgtgatgacatggat
ggagatggaaaacatctggacaactgcggccaaatttccaaatcgtgaccaacgg
gacaaggatggatgtgggtggggatgcctgtgacatgttgc
aaccagtctgtatgtggataatgatctgtggacttgcacatcaggacagt
gatggagatggcaccaggacacagacaactgcggccaccgttgcacatgtgcca
ctggacaccgataaggatggattttggcgttgcggatgtgtatgatgacatgtgg
atccccagacctgggtggccctggaccagacactgcggcgttgcggccaaacccagg
gaggatagcaacacgcgaaaggatggggatgcctgtgacatgttgc
gtcatcgatcgatgcgtctggccaggaaacgcggatgcacccgttgc
gcttaccagaccgtggccctggatcttgcggatggggatgcggccatgc
gttgcaccaggcatgggatgttgcggatgc
gggtacacagcttgcggatgttgcggatgc
gatgtgactatgcggatgc
atgtggaaaggcggacatggc
ctggcattcagtcaggatgc
tccctgtggc
aatgtggc
ggc
atagacaccc
atctgg
caaaaccc
tcgg
atata
aaaaaaa
aaaaaaa

protein:

MLAPRGAAVLLLHLVLQRWLAAGAQATPQVF DLLPSSSQRLNPGALLPVLTDPALNOLYV
ISTFKLQTKSSATIFGLYSSTDNSKYFEFTVMGRILSKAILRLYKNDGKVHLVVFNNLQLA
DGRHRILLRLSNLQRGAGSLELYLDCIQVDSVHNLPRAFAGPSQKPEIERTFQRKPO
DFLEELKLVVRGSLFQVASLQDCFLQQSEPLAATGTGDFNRQFLGQMTQLNQLLGEVKDL
LRQQVKETSFLRNNTIAECQACGPLKFQSPTPSTVAPAPPAPPTRPPRCDSNPCFRGVQ
CTDSRDGFQCGPCPEGYTGNGITCIDVDECKYHPCYPGVHICINLSPGFRCDACFVGF
MVQGVG15FAKSNKQVCTIDECRNGACVPNS1CVNTLGSYRCGPCKPGYTGDQIRGCKV
ERNCRNPELNPCSVNAQCIEERQGDVTCVCGVGWAGDGY1CGKDVIDSYDEELPCSAR
NCKKDNCVKYVPNSQGEDADRDGIGDACDEADGDGILNEQDNCVLIHNVDQRNSDKDIFG
DACDNCLSVLNNDQKDTDGDGRGDAACDDDMGDGDIKNILDNCPKFPNRDQRDKDGDGVGD

FIGURE 3B

ACDSCPDSNPNSDNDLVGDSCDTNQDSGDGHQDSTDNCPTVINSAQLDTDKGIG
DECDDDDDDNDGIPDLVPPGPDNCRLVPNPAQEDSNSDGVGDIICESDFDQQVIDRIDVCP
ENAEVTLTDFRAYQTVGLDPEGDAQIDPNWVVLNQGMEIVQTMNSDPGLAVGYTAFNGVD
FEGTFHVNTQTDDDYAGFIFGYQDSSSFYVVMWKQTEQTYWQATPFRAVAEPGIQLKAVK
SKTGPGEHLRNSLWHTGDTSDQVRLLWKDSRNVGWEDKVSYRWFLQHRPQVGYIRVRFYE
GSELVADSGVTIDTTMRGGRGLGVFCFSQENIIWSNLKYRCNDTIPEDFQEFTQNFDRFD
N

FIGURE 3C

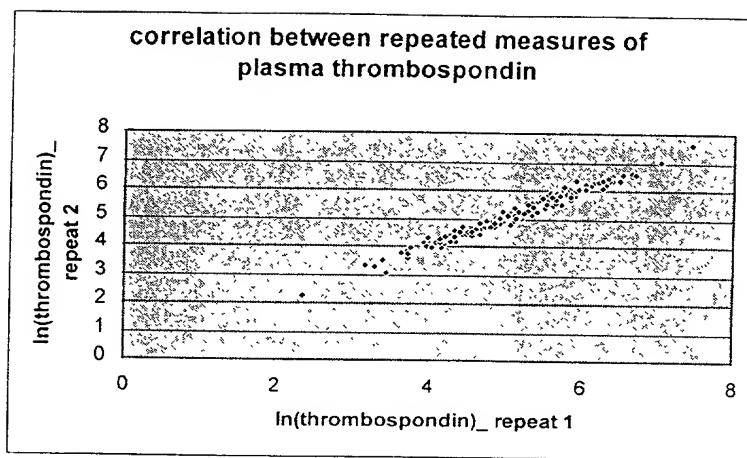


FIGURE 4